Genetics Home Reference

Your Guide to Understanding Genetic Conditions

Handbook

Help Me Understand Genetics

The Basics: Genes and How They Work

Reprinted from Genetics Home Reference (http://ghr.nlm.nih.gov/)

Lister Hill National Center for Biomedical Communications
U.S. National Library of Medicine
National Institutes of Health
Department of Health & Human Services

Published September 22, 2006

Chapter 1

The Basics: Genes and How They Work

Table of Contents

What is a cell?

Cells are the basic building blocks of all living things. The human body is composed of trillions of cells. They provide structure for the body, take in nutrients from food, convert those nutrients into energy, and carry out specialized functions. Cells also contain the body's hereditary material and can make copies of themselves.

Cells have many parts, each with a different function. Some of these parts, called organelles, are specialized structures that perform certain tasks within the cell. Human cells contain the following major parts, listed in alphabetical order:

Cytoplasm (illustration on page 4)

The cytoplasm is fluid inside the cell that surrounds the organelles.

Endoplasmic reticulum (ER) (illustration on page 5)

This organelle helps process molecules created by the cell and transport them to their specific destinations either inside or outside the cell.

Golgi apparatus (illustration on page 5)

The golgi apparatus packages molecules processed by the endoplasmic reticulum to be transported out of the cell.

Lysosomes and peroxisomes (illustration on page 5)

These organelles are the recycling center of the cell. They digest foreign bacteria that invade the cell, rid the cell of toxic substances, and recycle worn-out cell components.

Mitochondria (illustration on page 6)

Mitochondria are complex organelles that convert energy from food into a form that the cell can use. They have their own genetic material, separate from the DNA in the nucleus, and can make copies of themselves.

Nucleus (illustration on page 6)

The nucleus serves as the cell's command center, sending directions to the cell to grow, mature, divide, or die. It also houses DNA (deoxyribonucleic acid), the cell's hereditary material. The nucleus is surrounded by a membrane called the nuclear envelope, which protects the DNA and separates the nucleus from the rest of the cell.

Plasma membrane (illustration on page 6)

The plasma membrane is the outer lining of the cell. It separates the cell from its environment and allows materials to enter and leave the cell.

Ribosomes (illustration on page 7)

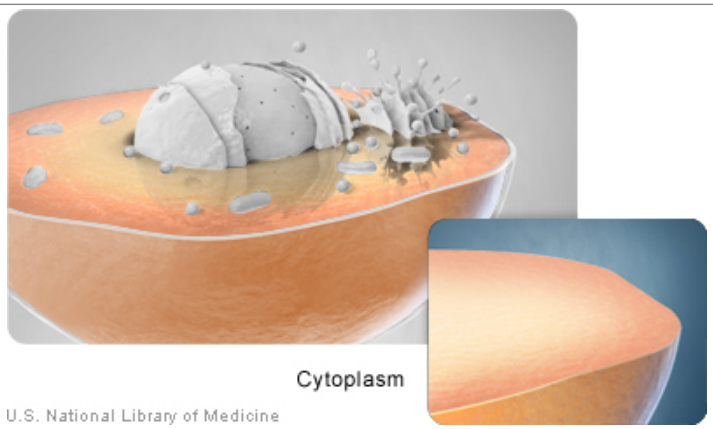
Ribosomes are organelles that process the cell's genetic instructions to create proteins. These organelles can float freely in the cytoplasm or be connected to the endoplasmic reticulum (see above).

For more information about cells:

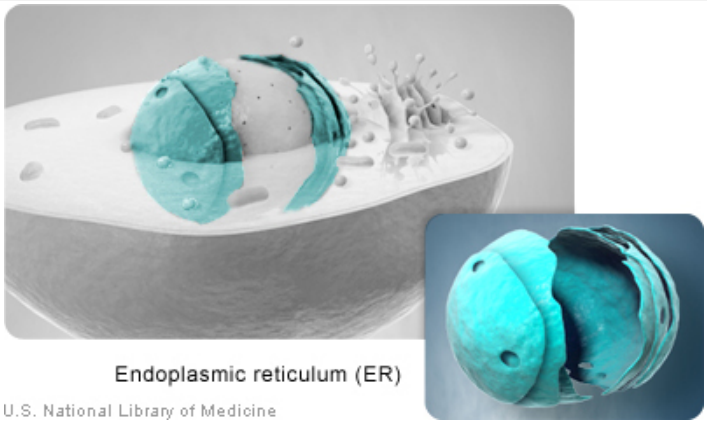
The NCBI Science Primer offers additional information about the structure and function of cells in the chapter titled What is a cell? (http://www.ncbi.nlm.nih.gov/About/primer/genetics_cell.html). Scroll down to the heading "Cell Structures: The Basics."

Mitochondria are essential organelles in cells. They have their own DNA, which is known as the mitochondrial genome. The Wellcome Trust provides a brief description (http://genome.wellcome.ac.uk/doc_wtd020740.html) of mitochondria and the mitochondrial genome.

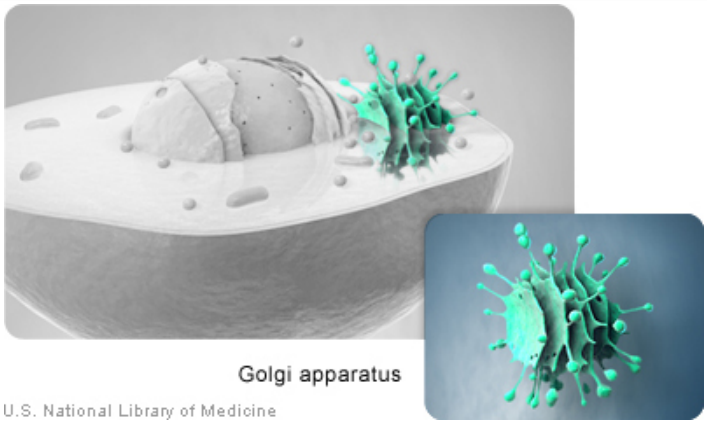
Illustrations



The cytoplasm surrounds the cell's nucleus and organelles.



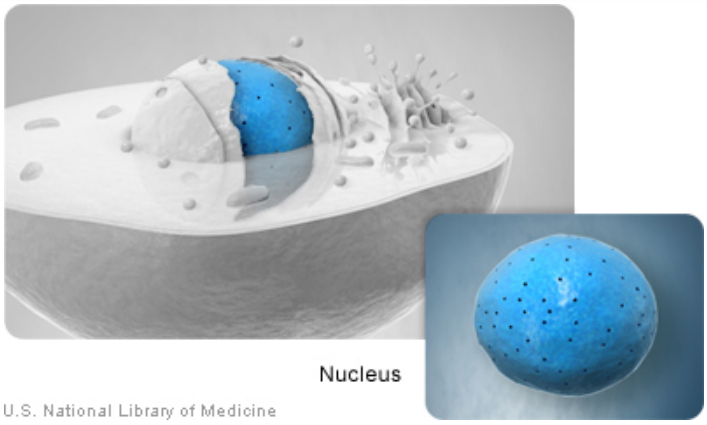
The endoplasmic reticulum is involved in molecule processing and transport.



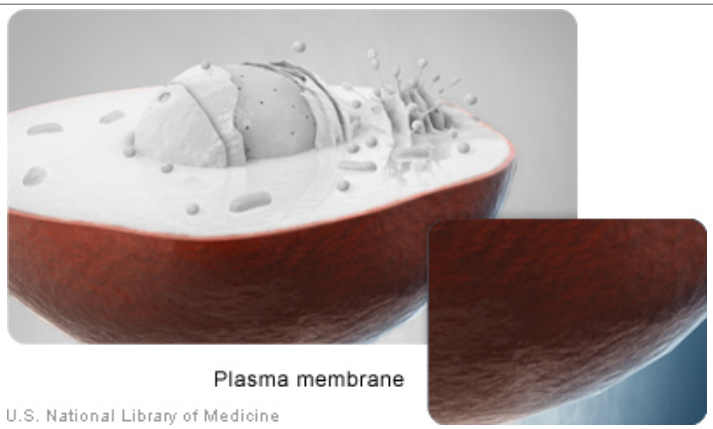
The Golgi apparatus is involved in packaging molecules for export from the cell.

Lysosomes and peroxisomes destroy toxic substances and recycle worn-out cell parts.

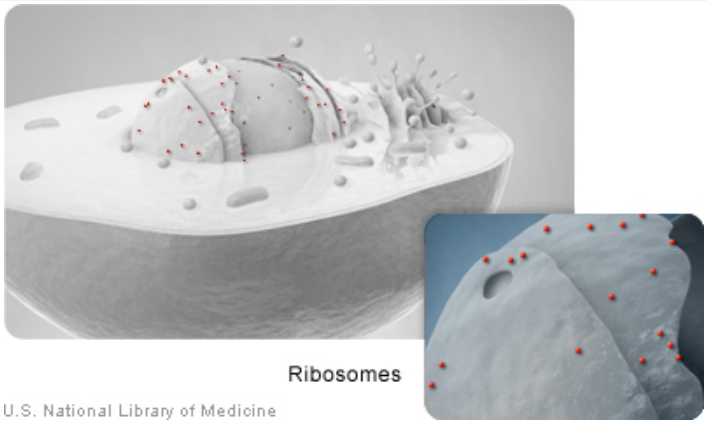
Mitochondria provide the cell's energy.



The nucleus contains most of the cell's genetic material.



The plasma membrane is the outer covering around the cell.



Ribosomes use the cell’s genetic instructions to make proteins.

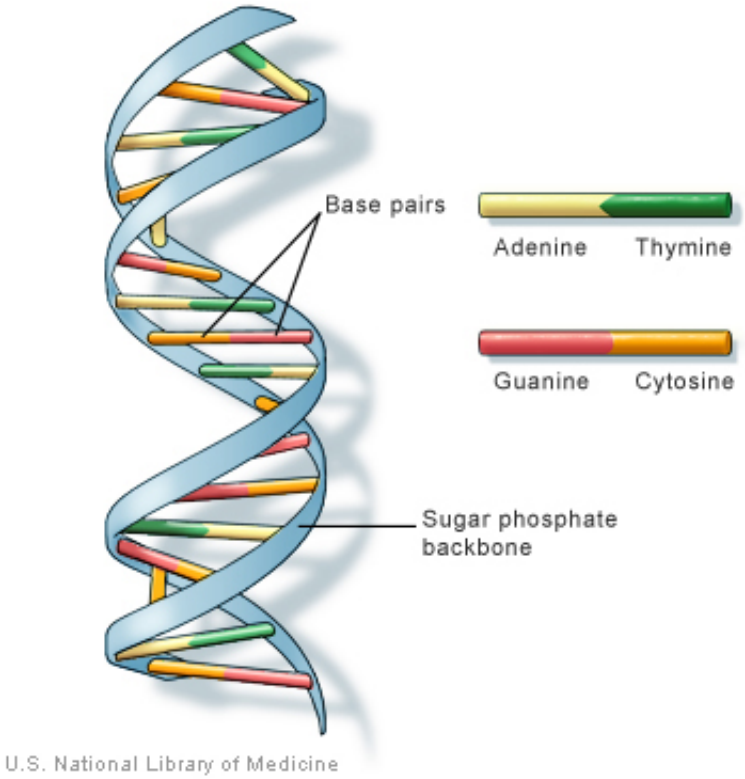
What is DNA?

DNA, or deoxyribonucleic acid, is the hereditary material in humans and almost all other organisms. Nearly every cell in a person's body has the same DNA. Most DNA is located in the cell nucleus (where it is called nuclear DNA), but a small amount of DNA can also be found in the mitochondria (where it is called mitochondrial DNA or mtDNA).

The information in DNA is stored as a code made up of four chemical bases: adenine (A), guanine (G), cytosine (C), and thymine (T). Human DNA consists of about 3 billion bases, and more than 99 percent of those bases are the same in all people. The order, or sequence, of these bases determines the information available for building and maintaining an organism, similar to the way in which letters of the alphabet appear in a certain order to form words and sentences.

DNA bases pair up with each other, A with T and C with G, to form units called base pairs. Each base is also attached to a sugar molecule and a phosphate molecule. Together, a base, sugar, and phosphate are called a nucleotide. Nucleotides are arranged in two long strands that form a spiral called a double helix. The structure of the double helix is somewhat like a ladder, with the base pairs forming the ladder's rungs and the sugar and phosphate molecules forming the vertical sidepieces of the ladder.

An important property of DNA is that it can replicate, or make copies of itself. Each strand of DNA in the double helix can serve as a pattern for duplicating the sequence of bases. This is critical when cells divide because each new cell needs to have an exact copy of the DNA present in the old cell.



DNA is a double helix formed by base pairs attached to a sugar-phosphate backbone.

For more information about DNA:

For additional information about the structure of DNA, please refer to the chapter called What Is A Genome? (http://www.ncbi.nih.gov/About/primer/genetics_genome.html) in the NCBI Science Primer. Scroll down to the heading “The Physical Structure of the Human Genome.”

What is mitochondrial DNA?

Although most DNA is packaged in chromosomes within the nucleus, mitochondria also have a small amount of their own DNA. This genetic material is known as mitochondrial DNA or mtDNA.

Mitochondria (illustration on page 6) are structures within cells that convert the energy from food into a form that cells can use. Each cell contains hundreds to thousands of mitochondria, which are located in the fluid that surrounds the nucleus (the cytoplasm).

Mitochondria produce energy through a process called oxidative phosphorylation. This process uses oxygen and simple sugars to create adenosine triphosphate (ATP), the cell's main energy source. A set of enzyme complexes, designated as complexes I-V, carry out oxidative phosphorylation within mitochondria.

In addition to energy production, mitochondria play a role in several other cellular activities. For example, mitochondria help regulate the self-destruction of cells (apoptosis). They are also necessary for the production of substances such as cholesterol and heme (a component of hemoglobin, the molecule that carries oxygen in the blood).

Mitochondrial DNA contains 37 genes, all of which are essential for normal mitochondrial function. Thirteen of these genes provide instructions for making enzymes involved in oxidative phosphorylation. The remaining genes provide instructions for making molecules called transfer RNAs (tRNAs) and ribosomal RNAs (rRNAs), which are chemical cousins of DNA. These types of RNA help assemble protein building blocks (amino acids) into functioning proteins.

For more information about mitochondria and mitochondrial DNA:

Molecular Expressions, a web site from the Florida State University Research Foundation, offers an illustrated introduction to mitochondria and mitochondrial DNA (http://micro.magnet.fsu.edu/cells/mitochondria/mitochondria.html).

An overview of mitochondria (http://www.neuro.wustl.edu/neuromuscular/mitosyn.html#mitogen), including information about mitochondrial DNA, is available from the Neuromuscular Disease Center at Washington University.

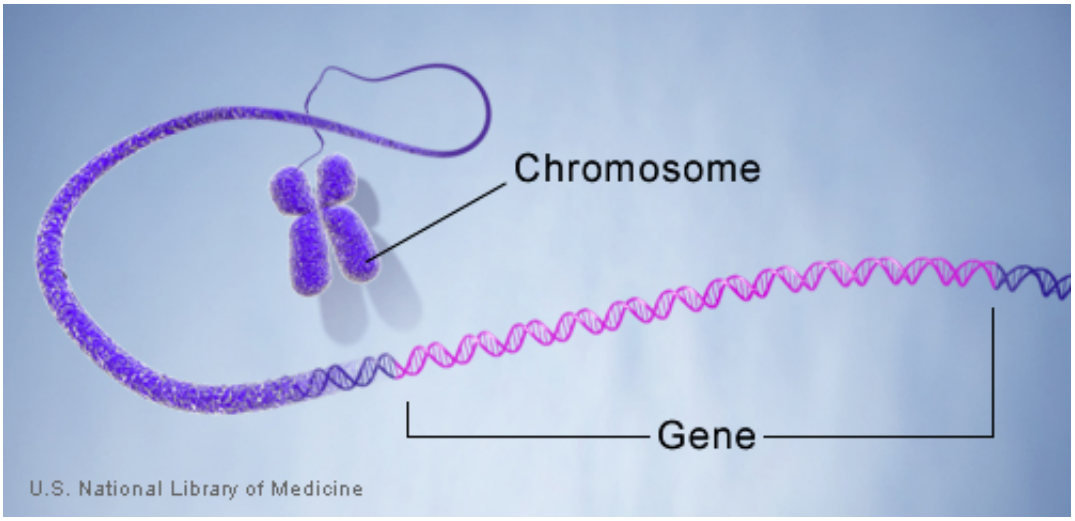
The Wellcome Trust provides a discussion of the mitochondrial genome (http://genome.wellcome.ac.uk/doc_WTD020740.html), including an explanation of why mitochondria have their own DNA.

The Howard Hughes Research Institute offers an article about recent research into mitochondrial function (http://www.hhmi.org/bulletin/may2006/features/mitochondria.html).

What is a gene?

A gene is the basic physical and functional unit of heredity. Genes, which are made up of DNA, act as instructions to make molecules called proteins. In humans, genes vary in size from a few hundred DNA bases to more than 2 million bases. The Human Genome Project has estimated that humans have between 20,000 and 25,000 genes.

Every person has two copies of each gene, one inherited from each parent. Most genes are the same in all people, but a small number of genes (less than 1 percent of the total) are slightly different between people. Alleles are forms of the same gene with small differences in their sequence of DNA bases. These small differences contribute to each person's unique physical features.



Genes are made up of DNA. Each chromosome contains many genes.

For more information about genes:

Genetics Home Reference provides consumer-friendly gene summaries (http://ghr.nlm.nih.gov/ghr/genesBySymbol) that include an explanation of each gene's normal function and how mutations in the gene cause particular genetic conditions.

The National Institute of General Medical Sciences offers additional information about DNA and genes in its publication Genetic Basics. Refer to the publication's introduction, A Science Called Genetics (http://publications.nigms.nih.gov/genetics/science.html).

The Wellcome Trust provides a basic overview of gene structure (http://genome.wellcome.ac.uk/doc_WTD020755.html).

The Centre for Genetics Education offers a fact sheet that introduces genes and chromosomes (http://www.genetics.com.au/factsheet/01.htm).

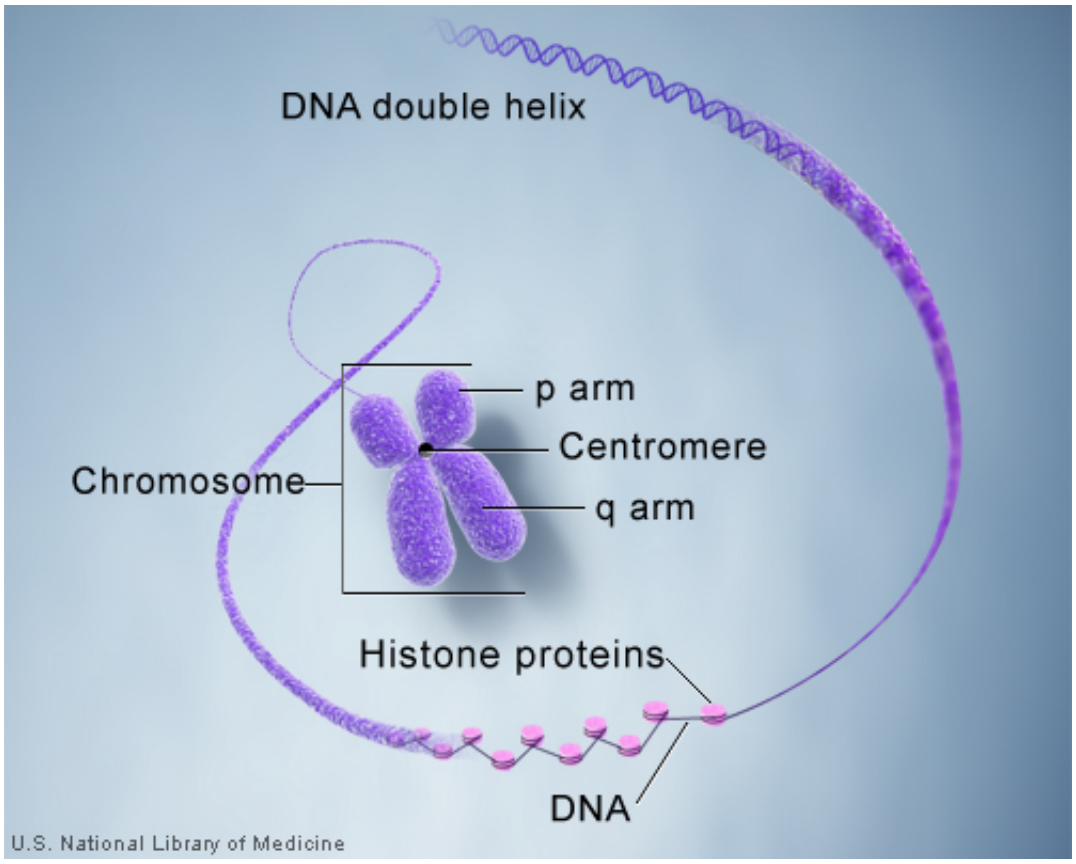
For more information about genes, refer to the chapter titled What is a Genome? (http://www.ncbi.nih.gov/About/primer/genetics_genome.html) in the NCBI Science Primer.

What is a chromosome?

In the nucleus of each cell, the DNA molecule is packaged into thread-like structures called chromosomes. Each chromosome is made up of DNA tightly coiled many times around proteins called histones that support its structure.

Chromosomes are not visible in the cell's nucleus—not even under a microscope—when the cell is not dividing. However, the DNA that makes up chromosomes becomes more tightly packed during cell division and is then visible under a microscope. Most of what researchers know about chromosomes was learned by observing chromosomes during cell division.

Each chromosome has a constriction point called the centromere, which divides the chromosome into two sections, or "arms." The short arm of the chromosome is labeled the "p arm." The long arm of the chromosome is labeled the "q arm." The location of the centromere on each chromosome gives the chromosome its characteristic shape, and can be used to help describe the location of specific genes.



DNA and histone proteins are packaged into structures called chromosomes.

For more information about chromosomes:

Genetics Home Reference provides information about each human chromosome (http://ghr.nlm.nih.gov/ghr/chromosomes) written in lay language.

The Centre for Genetics Education offers a fact sheet that introduces genes and chromosomes (http://www.genetics.com.au/factsheet/01.htm).

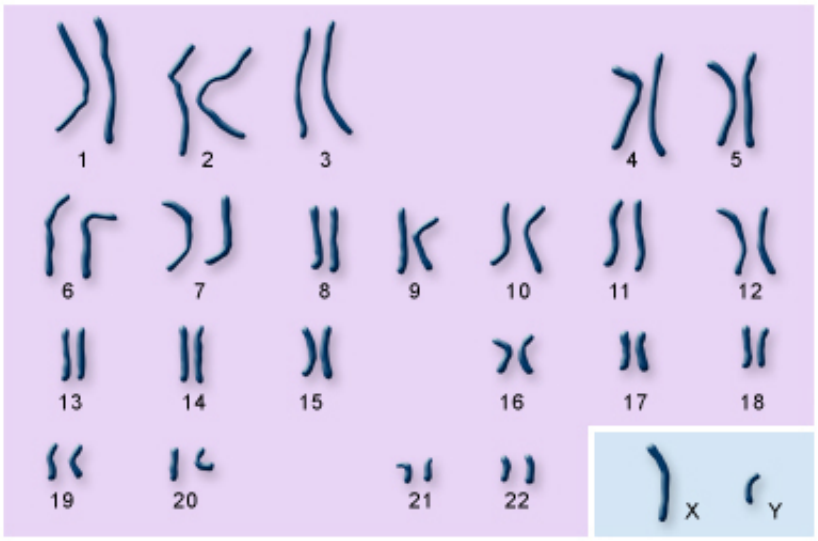
The NCBI Science Primer includes a discussion of the DNA that makes up chromosomes in the chapter called What Is A Genome? (http://www.ncbi.nih.gov/About/primer/genetics_genome.html). Scroll down to the heading “Structural Genes, Junk DNA and Regulatory Sequences.”

The U.S. Department of Energy Office of Science offers a list of Chromosome FAQs (http://www.ornl.gov/sci/techresources/Human_Genome/posters/chromosome/faqs.shtml).

Information about centromeres and their role in cell division (http://genome.wellcome.ac.uk/doc_wtd020744.html) is available from the Wellcome Trust.

How many chromosomes do people have?

In humans, each cell normally contains 23 pairs of chromosomes, for a total of 46. Twenty-two of these pairs, called autosomes, look the same in both males and females. The 23rd pair, the sex chromosomes, differ between males and females. Females have two copies of the X chromosome, while males have one X and one Y chromosome.



The 22 autosomes are numbered by size. The other two chromosomes, X and Y, are the sex chromosomes. This picture of the human chromosomes lined up in pairs is called a karyotype.

For more information about the 23 pairs of human chromosomes:

Genetics Home Reference provides information about each human chromosome (http://ghr.nlm.nih.gov/ghr/chromosomes) written in lay language.

How do geneticists indicate the location of a gene?

Geneticists use maps to describe the location of a particular gene on a chromosome. One type of map uses the cytogenetic location to describe a gene's position. The cytogenetic location is based on a distinctive pattern of bands created when chromosomes are stained with certain chemicals. Another type of map uses the molecular location, a precise description of a gene's position on a chromosome. The molecular location is based on the sequence of DNA building blocks (base pairs) that make up the chromosome.

Cytogenetic location

Geneticists use a standardized way of describing a gene's cytogenetic location. In most cases, the location describes the position of a particular band on a stained chromosome:

17q12

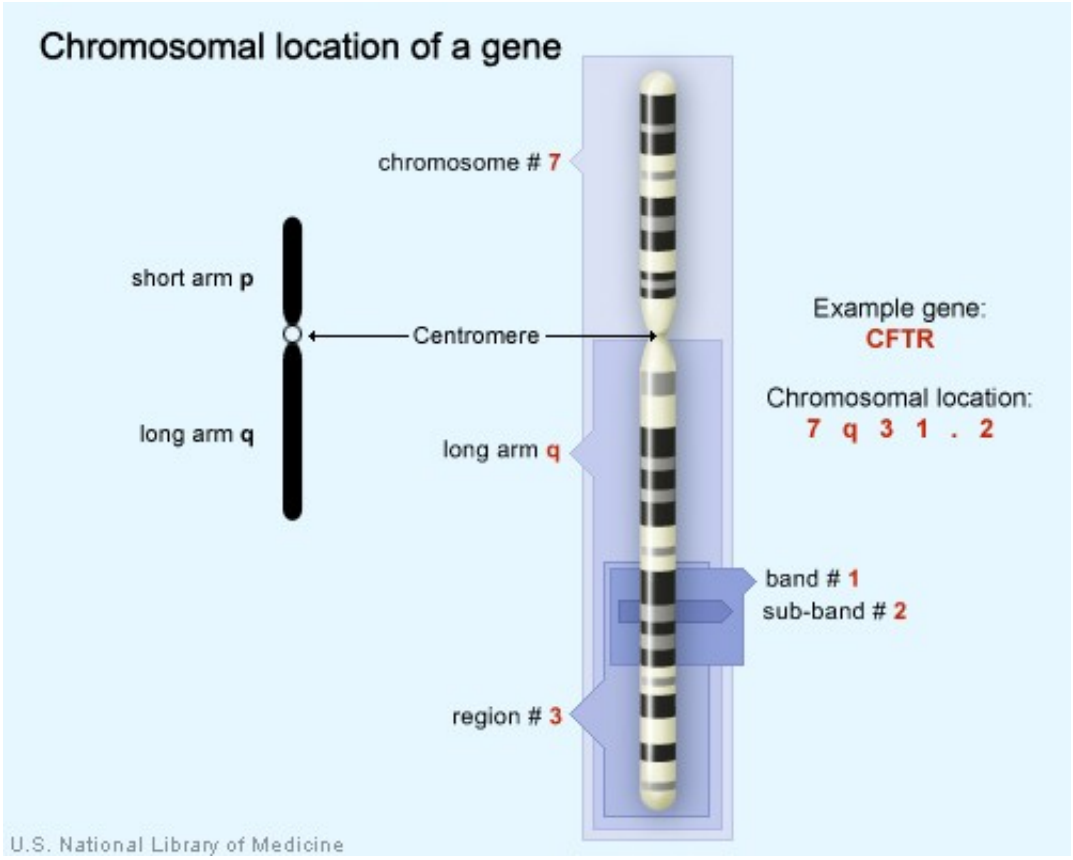
It can also be written as a range of bands, if less is known about the exact location:

17q12-q21

The combination of numbers and letters provide a gene's "address" on a chromosome. This address is made up of several parts:

- The chromosome on which the gene can be found. The first number or letter used to describe a gene's location represents the chromosome. Chromosomes 1 through 22 (the autosomes) are designated by their chromosome number. The sex chromosomes are designated by X or Y.
- The arm of the chromosome. Each chromosome is divided into two sections (arms) based on the location of a narrowing (constriction) called the centromere. By convention, the shorter arm is called p, and the longer arm is called q. The chromosome arm is the second part of the gene's address. For example, 5q is the long arm of chromosome 5, and Xp is the short arm of the X chromosome.
- The position of the gene on the p or q arm. The position of a gene is based on a distinctive pattern of light and dark bands that appear when the chromosome is stained in a certain way. The position is usually designated by two digits (representing a region and a band), which are sometimes followed by a decimal point and one or more additional digits (representing sub-bands within a light or dark area). The number indicating the gene position increases with distance from the centromere. For example: 14q21 represents position 21 on the long arm of chromosome 14. 14q21 is closer to the centromere than 14q22.

Sometimes, the abbreviations “cen” or “ter” are also used to describe a gene’s cytogenetic location. “Cen” indicates that the gene is very close to the centromere. For example, 16pcen refers to the short arm of chromosome 16 near the centromere. “Ter” stands for terminus, which indicates that the gene is very close to the end of the p or q arm. For example, 14qter refers to the tip of the long arm of chromosome 14. (“Tel” is also sometimes used to describe a gene’s location. “Tel” stands for telomeres, which are at the ends of each chromosome. The abbreviations “tel” and “ter” refer to the same location.)



The CFTR gene is located on the long arm of chromosome 7 at position 7q31.2.

Molecular location

The Human Genome Project, an international research effort completed in 2003, determined the sequence of base pairs for each human chromosome. This sequence information allows researchers to provide a more specific address than the cytogenetic location for many genes. A gene’s molecular address pinpoints the location of that gene in terms of base pairs. For example, the molecular location of the APOE gene on chromosome 19 begins with base pair 50,100,901 and ends with base pair 50,104,488. This range describes the gene’s precise position on

chromosome 19 and indicates the size of the gene (3,588 base pairs). Knowing a gene's molecular location also allows researchers to determine exactly how far the gene is from other genes on the same chromosome.

Different groups of researchers often present slightly different values for a gene's molecular location. Researchers interpret the sequence of the human genome using a variety of methods, which can result in small differences in a gene's molecular address. For example, the National Center for Biotechnology Information (NCBI) identifies the molecular location of the APOE gene as base pair 50,100,901 to base pair 50,104,488 on chromosome 19. The Ensembl database identifies the location of this gene as base pair 50,100,879 to base pair 50,104,489 on chromosome 19. Neither of these addresses is incorrect; they represent different interpretations of the same data. For consistency, Genetics Home Reference presents data from NCBI for the molecular location of genes.

For more information on genetic mapping:

The National Human Genome Research Institute explains how researchers create a genetic map (http://www.genome.gov/10000715).

The University of Washington provides a Cytogenetics Gallery (http://www.pathology.washington.edu/galleries/Cytogallery/main.php?file=intro) that includes a description of chromosome banding patterns (http://www.pathology.washington.edu/galleries/Cytogallery/main.php?file=banding+patterns).

The NCBI Science Primer offers additional detailed information about genome mapping (http://www.ncbi.nlm.nih.gov/About/primer/mapping.html).

What are proteins and what do they do?

Proteins are large, complex molecules that play many critical roles in the body. They do most of the work in cells and are required for the structure, function, and regulation of the body's tissues and organs.

Proteins are made up of hundreds or thousands of smaller units called amino acids, which are attached to one another in long chains. There are 20 different types of amino acids that can be combined to make a protein. The sequence of amino acids determines each protein's unique 3-dimensional structure and its specific function.

Proteins can be described according to their large range of functions in the body, listed in alphabetical order:

Examples of protein functions

Function	Description	Example
Antibody	Antibodies bind to specific foreign particles, such as viruses and bacteria, to help protect the body.	Immunoglobulin G (IgG) (illustration on page 20)
Enzyme	Enzymes carry out almost all of the thousands of chemical reactions that take place in cells. They also assist with the formation of new molecules by reading the genetic information stored in DNA.	Phenylalanine hydroxylase (illustration on page 21)
Messenger	Messenger proteins, such as some types of hormones, transmit signals to coordinate biological processes between different cells, tissues, and organs.	Growth hormone (illustration on page 22)
Structural component	These proteins provide structure and support for cells. On a larger scale, they also allow the body to move.	Actin (illustration on page 23)
Transport/storage	These proteins bind and carry atoms and small molecules within cells and throughout the body.	Ferritin (illustration on page 24)

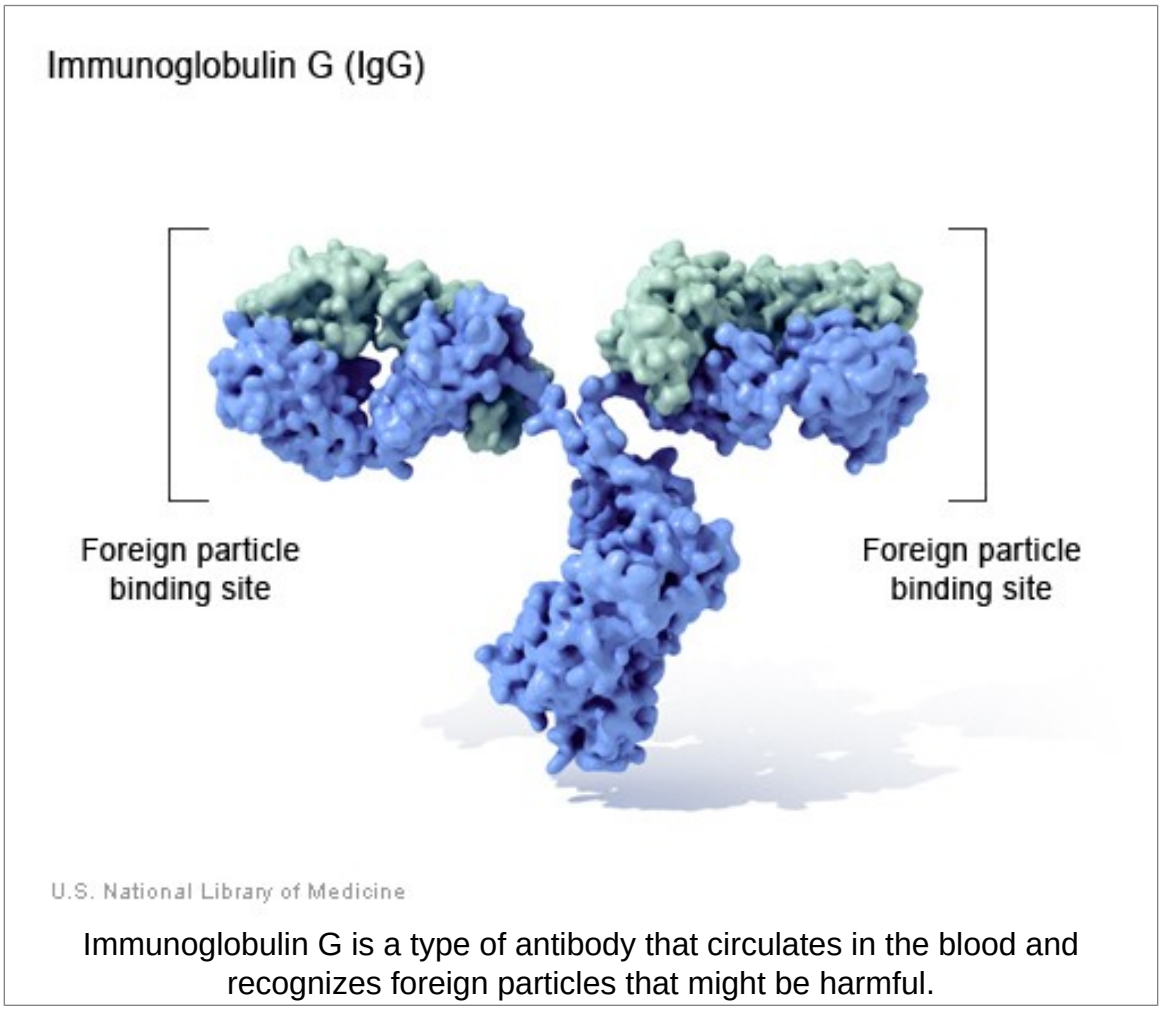
For more information about proteins and their functions:

Information about proteins and what they do is available from the National Institute of General Medical Sciences publication Genetic Basics. Refer to the section called From Genes to Proteins (http://publications.nigms.nih.gov/genetics/chapter1.html#a1).

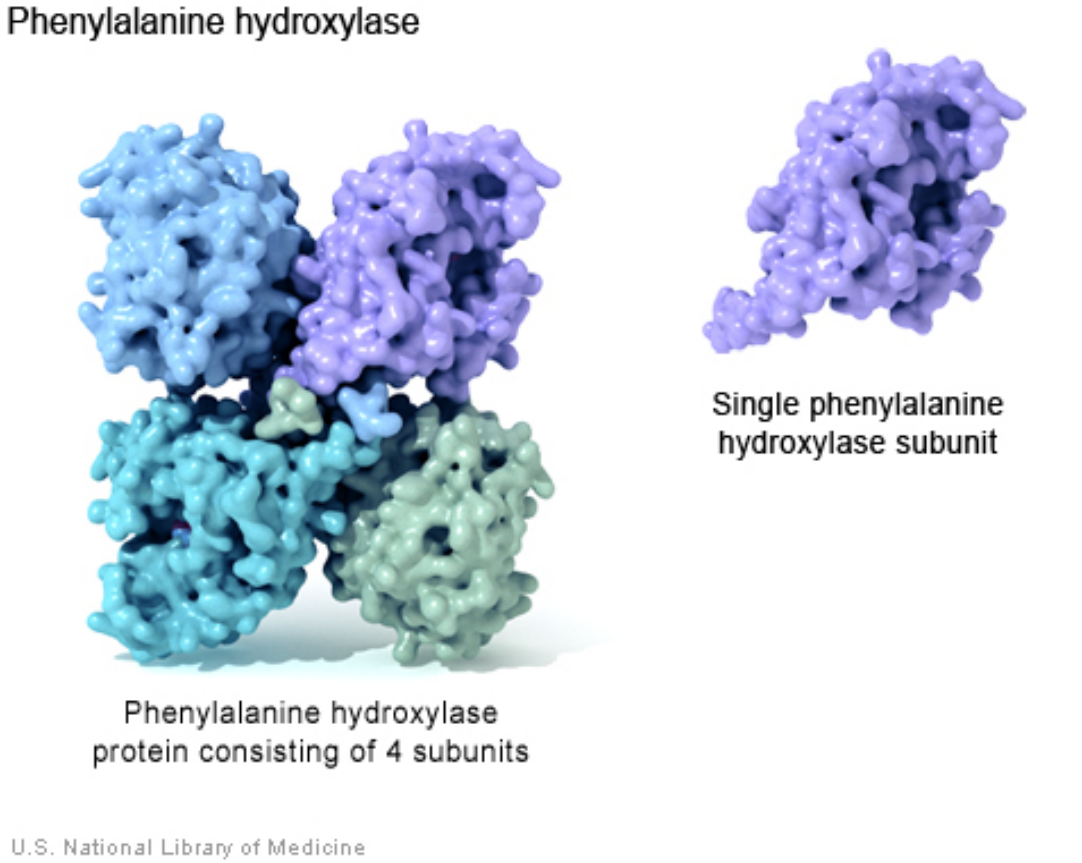
Additional discussion of the role of proteins can be found in the NCBI Science Primer in the chapter called What Is A Genome? (http://www.ncbi.nlm.nih.gov/About/primer/genetics_genome.html). Scroll down to the heading “Proteins.”

The Wellcome Trust provides basic information about protein structure and function (http://genome.wellcome.ac.uk/doc_WTD020766.html).

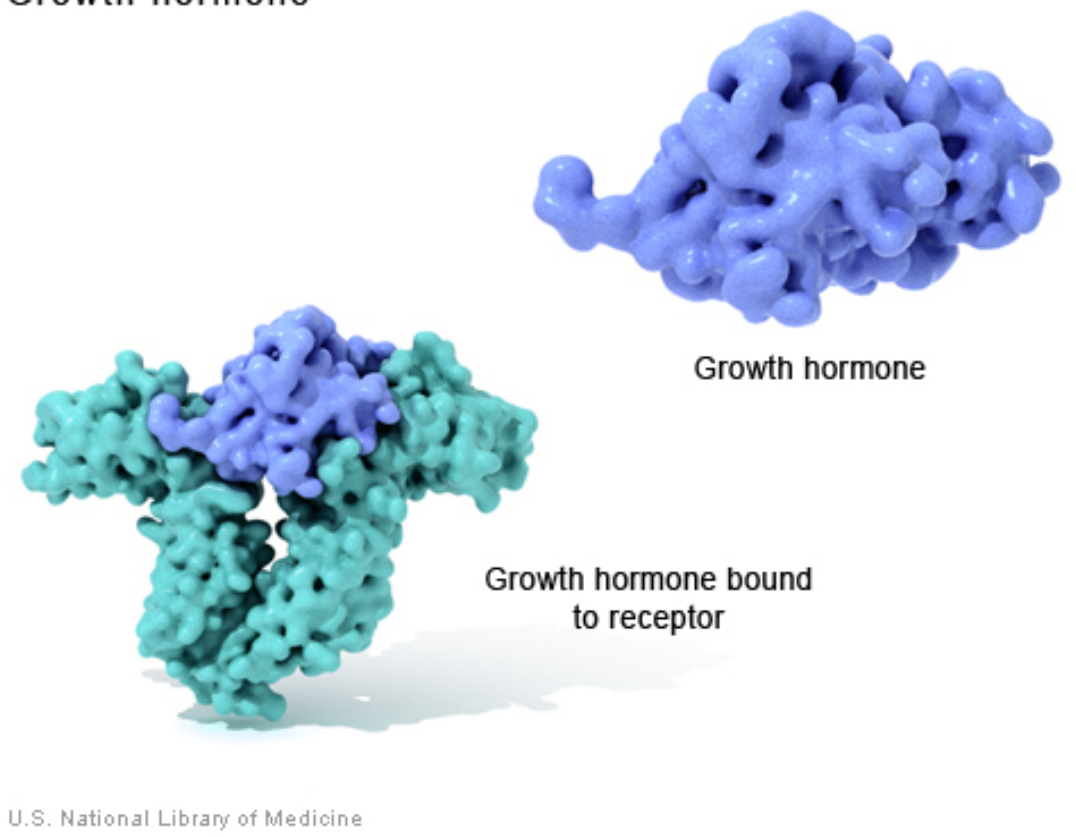
Illustrations



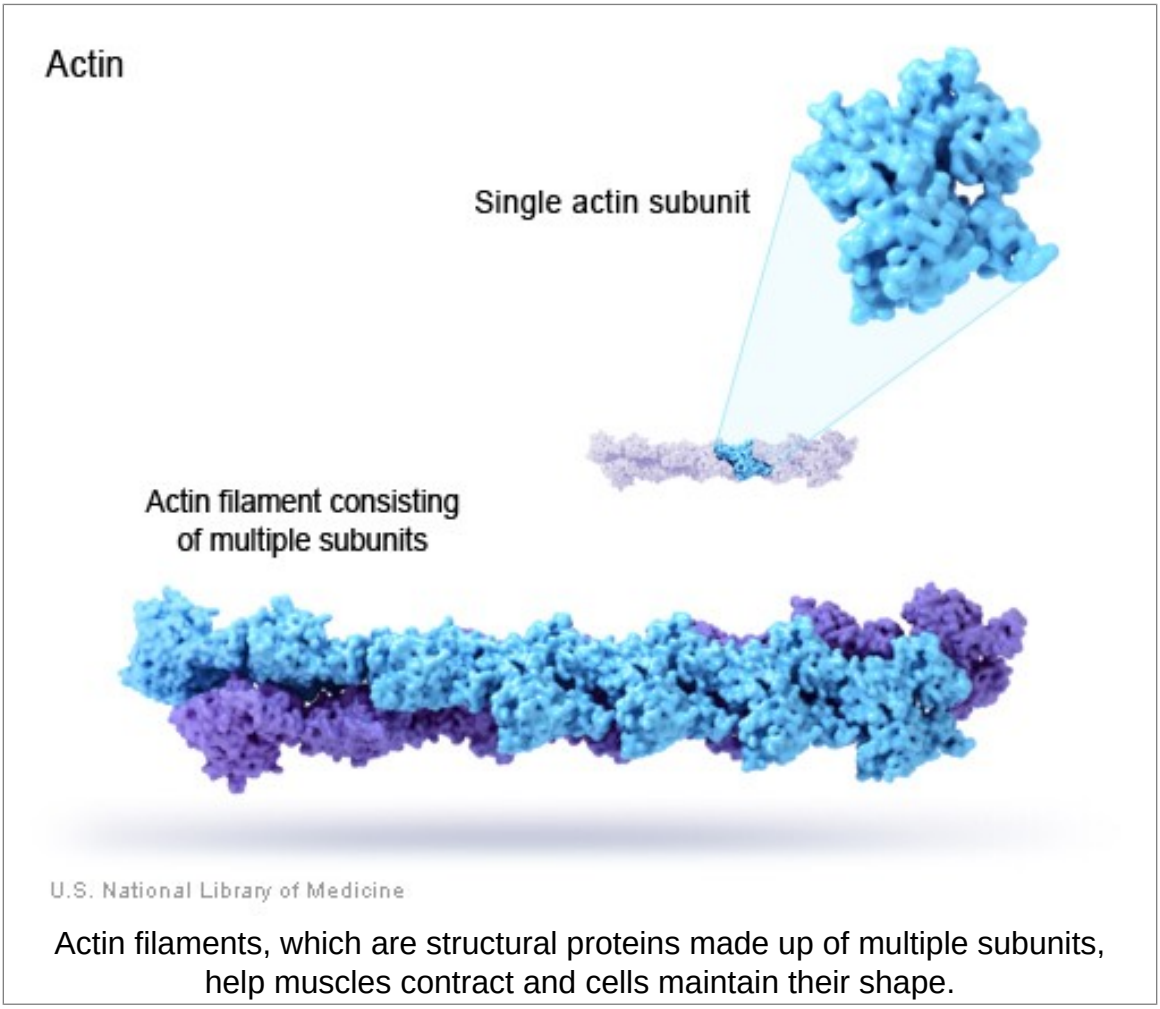
Immunoglobulin G is a type of antibody that circulates in the blood and recognizes foreign particles that might be harmful.



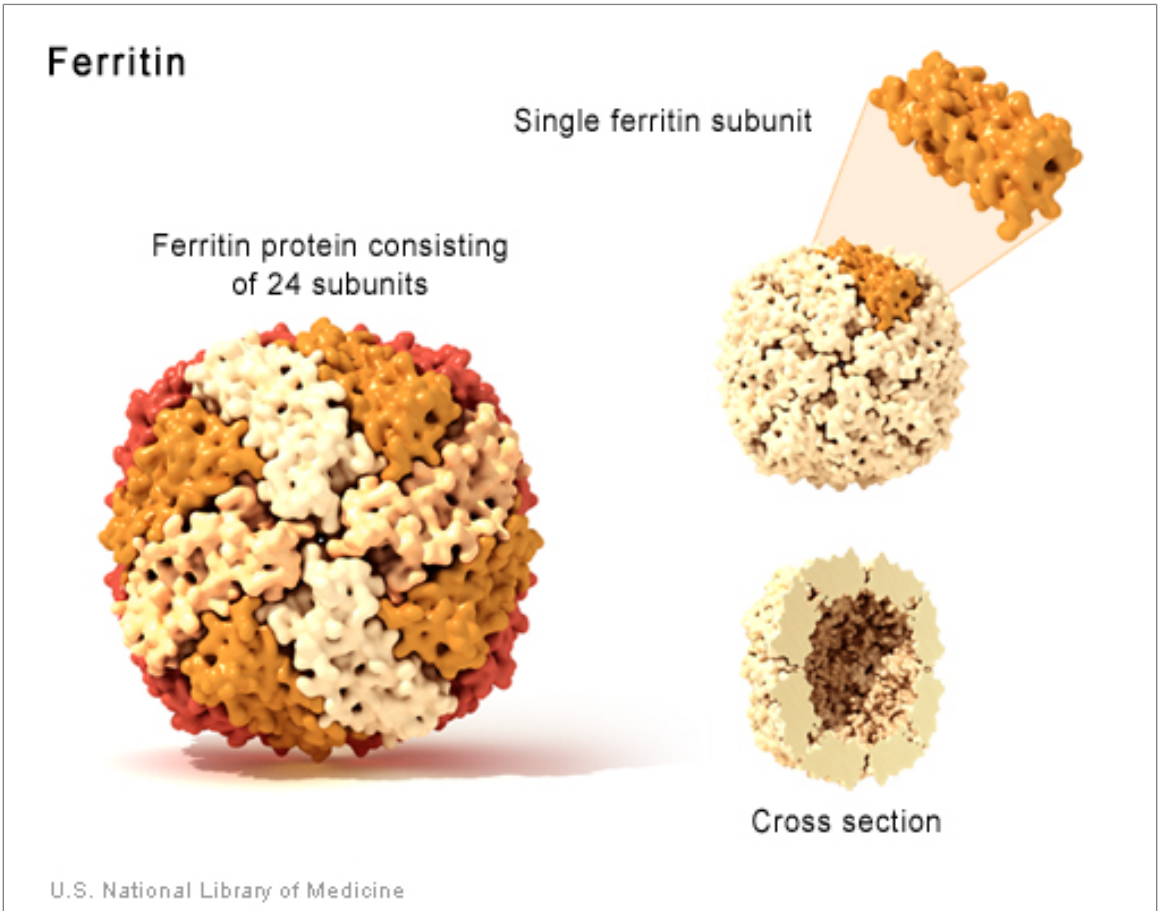
The functional phenylalanine hydroxylase enzyme is made up of four identical subunits. The enzyme converts the amino acid phenylalanine to another amino acid, tyrosine.



Growth hormone is a messenger protein made by the pituitary gland. It regulates cell growth by binding to a protein called a growth hormone receptor.



Actin filaments, which are structural proteins made up of multiple subunits, help muscles contract and cells maintain their shape.



Ferritin, a protein made up of 24 identical subunits, is involved in iron storage.

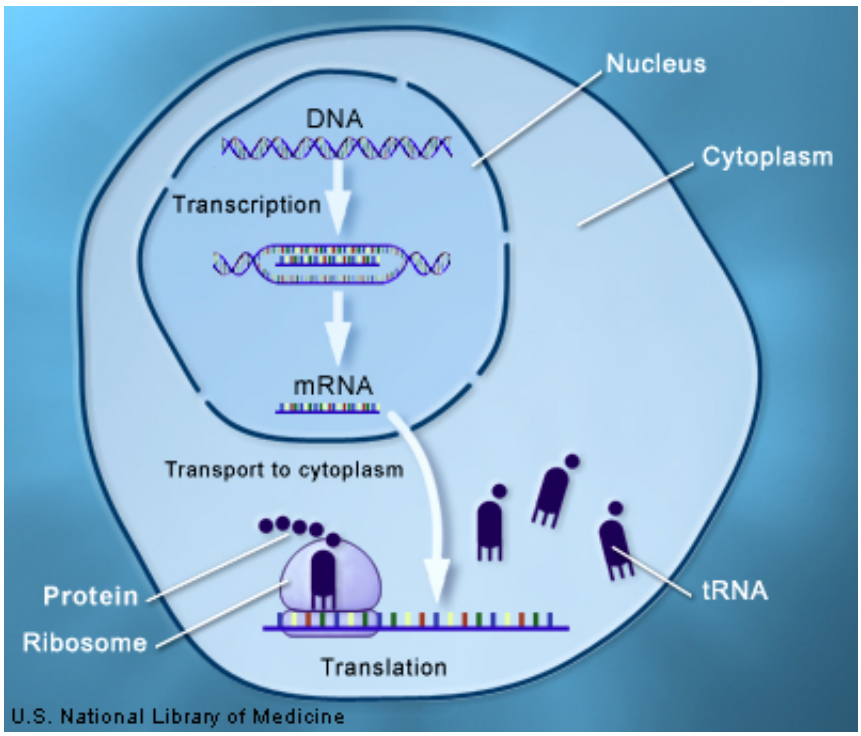
How does a gene make a protein?

Most genes contain the information needed to make functional molecules called proteins. (A few genes produce other molecules that help the cell assemble proteins.) The journey from gene to protein is complex and tightly controlled within each cell. It consists of two major steps: transcription and translation. Together, transcription and translation are known as gene expression.

During the process of transcription, the information stored in a gene's DNA is transferred to a similar molecule called RNA (ribonucleic acid) in the cell nucleus. Both RNA and DNA are made up of a chain of nucleotide bases, but they have slightly different chemical properties. The type of RNA that contains the information for making a protein is called messenger RNA (mRNA) because it carries the information, or message, from the DNA out of the nucleus into the cytoplasm.

Translation, the second step in getting from a gene to a protein, takes place in the cytoplasm. The mRNA interacts with a specialized complex called a ribosome, which "reads" the sequence of mRNA bases. Each sequence of three bases, called a codon, usually codes for one particular amino acid. (Amino acids are the building blocks of proteins.) A type of RNA called transfer RNA (tRNA) assembles the protein, one amino acid at a time. Protein assembly continues until the ribosome encounters a "stop" codon (a sequence of three bases that does not code for an amino acid).

The flow of information from DNA to RNA to proteins is one of the fundamental principles of molecular biology. It is so important that it is sometimes called the "central dogma."



Through the processes of transcription and translation, information from genes is used to make proteins.

For more information about making proteins:

The Wellcome Trust provides a brief overview of gene expression (http://genome.wellcome.ac.uk/doc_WTD020757.html).

For a more detailed description of transcription and translation, refer to the NCBI Science Primer's chapter titled What Is A Genome? (http://www.ncbi.nlm.nih.gov/About/primer/genetics_genome.html). Scroll down to the heading "From Genes to Proteins: Start to Finish."

Can genes be turned on and off in cells?

Each cell expresses, or turns on, only a fraction of its genes. The rest of the genes are repressed, or turned off. The process of turning genes on and off is known as gene regulation. Gene regulation is an important part of normal development. Genes are turned on and off in different patterns during development to make a brain cell look and act different from a liver cell or a muscle cell, for example. Gene regulation also allows cells to react quickly to changes in their environments. Although we know that the regulation of genes is critical for life, this complex process is not yet fully understood.

Gene regulation can occur at any point during gene expression, but most commonly occurs at the level of transcription (when the information in a gene's DNA is transferred to mRNA). Signals from the environment or from other cells activate proteins called transcription factors. These proteins bind to regulatory regions of a gene and increase or decrease the level of transcription. By controlling the level of transcription, this process can determine the amount of protein product that is made by a gene at any given time.

For more information about gene regulation:

More information about gene regulation can be found in the NCBI Science Primer. Refer to the chapter called What Is A Genome? (http://www.ncbi.nlm.nih.gov/About/primer/genetics_genome.html) and scroll down to the headings "Gene Switching: Turning Genes On and Off," "Controlling Transcription," and "Controlling Translation."

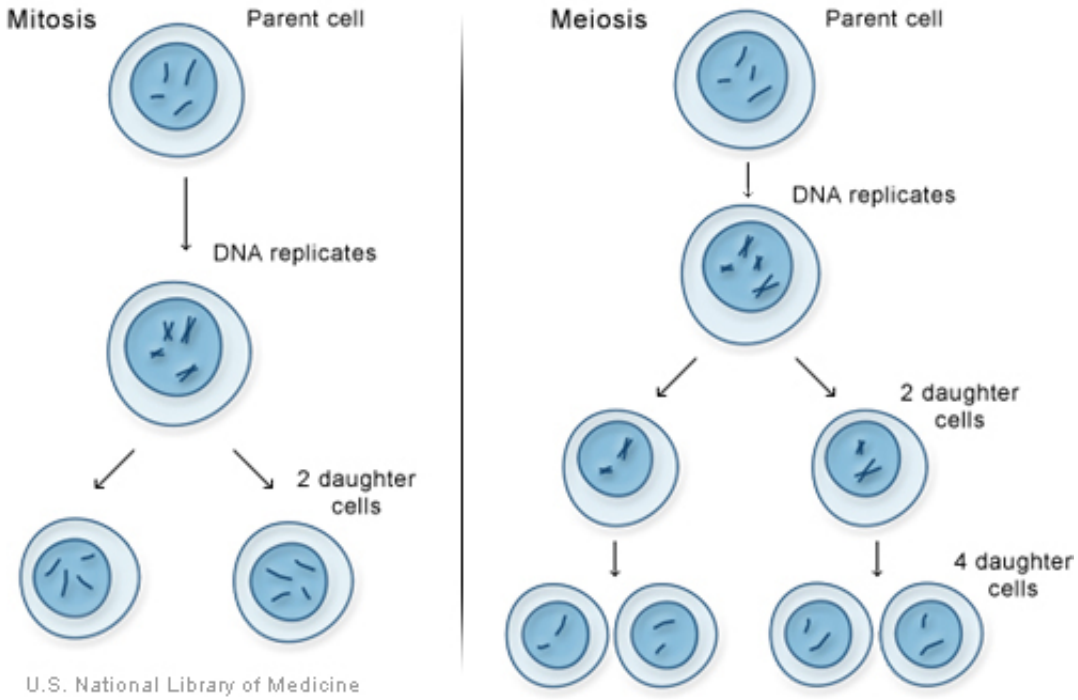
The National Institute of General Medical Science publication Genetic Basics also offers a discussion of gene regulation. Refer to the section called Controlling Genes (http://publications.nigms.nih.gov/genetics/chapter1.html#a3).

How do cells divide?

There are two types of cell division: mitosis and meiosis. Most of the time when people refer to “cell division,” they mean mitosis, the process of making new body cells. Meiosis is the type of cell division that creates egg and sperm cells.

Mitosis is a fundamental process for life. During mitosis, a cell duplicates all of its contents, including its chromosomes, and splits to form two identical daughter cells. Because this process is so critical, the steps of mitosis are carefully controlled by a number of genes. When mitosis is not regulated correctly, health problems such as cancer can result.

The other type of cell division, meiosis, ensures that humans have the same number of chromosomes in each generation. It is a two-step process that reduces the chromosome number by half—from 46 to 23—to form sperm and egg cells. When the sperm and egg cells unite at conception, each contributes 23 chromosomes so the resulting embryo will have the usual 46. Meiosis also allows genetic variation through a process of DNA shuffling while the cells are dividing.



Mitosis and meiosis, the two types of cell division.

For more information about cell division:

For a detailed summary of mitosis and meiosis, please refer to the chapter titled What Is A Cell? (http://www.ncbi.nlm.nih.gov/About/primer/genetics_cell.html) In the NCBI Science Primer. Scroll down to the heading “Making New Cells and Cell Types.”

How do genes control the growth and division of cells?

A variety of genes are involved in the control of cell growth and division. The cell cycle is the cell's way of replicating itself in an organized, step-by-step fashion. Tight regulation of this process ensures that a dividing cell's DNA is copied properly, any errors in the DNA are repaired, and each daughter cell receives a full set of chromosomes. The cycle has checkpoints (also called restriction points), which allow certain genes to check for mistakes and halt the cycle for repairs if something goes wrong.

If a cell has an error in its DNA that cannot be repaired, it may undergo programmed cell death (apoptosis) (illustration on page 31). Apoptosis is a common process throughout life that helps the body get rid of cells it doesn't need. Cells that undergo apoptosis break apart and are recycled by a type of white blood cell called a macrophage (illustration on page 31). Apoptosis protects the body by removing genetically damaged cells that could lead to cancer, and it plays an important role in the development of the embryo and the maintenance of adult tissues.

Cancer results from a disruption of the normal regulation of the cell cycle. When the cycle proceeds without control, cells can divide without order and accumulate genetic defects that can lead to a cancerous tumor (illustration on page 32).

For more information about cell growth and division:

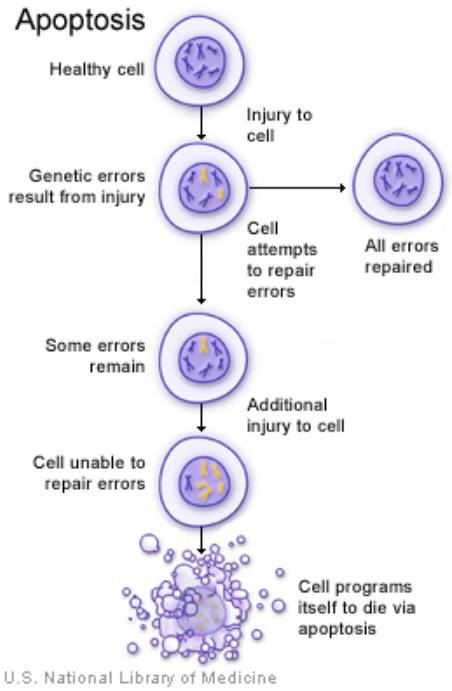
The National Institutes of Health's Apoptosis Interest Group (http://www.nih.gov/sigs/aig/Aboutapo.html) provides an introduction to programmed cell death.

Additional information about apoptosis is available from the National Institute of General Medical Sciences publication Genetic Basics. Refer to the section called Programmed Cell Death (http://publications.nigms.nih.gov/genetics/chapter3.html#c2).

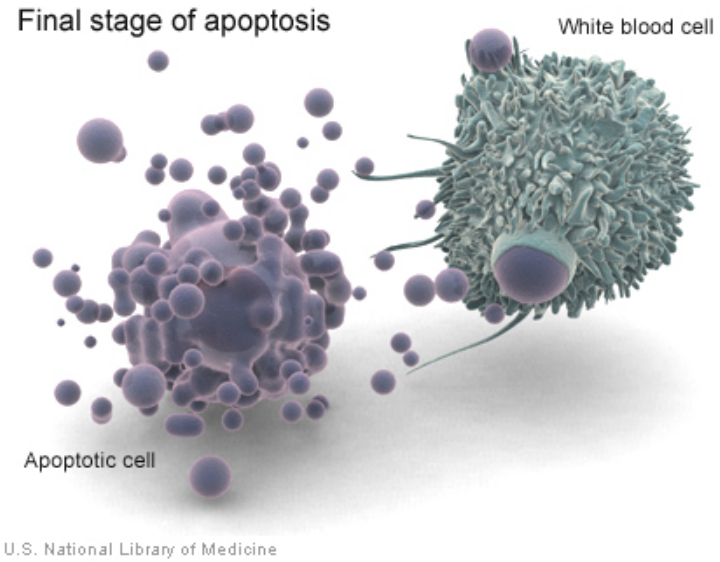
The National Cancer Institute offers several publications that explain the growth of cancerous tumors. These include What You Need To Know About Cancer—An Overview (http://www.cancer.gov/cancerinfo/wyntk/overview) and Understanding Cancer (http://www.cancer.gov/cancertopics/understandingcancer/cancer).

The Wellcome Trust offers brief overviews of the role of genes in cancer (http://genome.wellcome.ac.uk/doc_WTD020845.html) and the types of genes involved in cancer (http://genome.wellcome.ac.uk/doc_WTD020846.html).

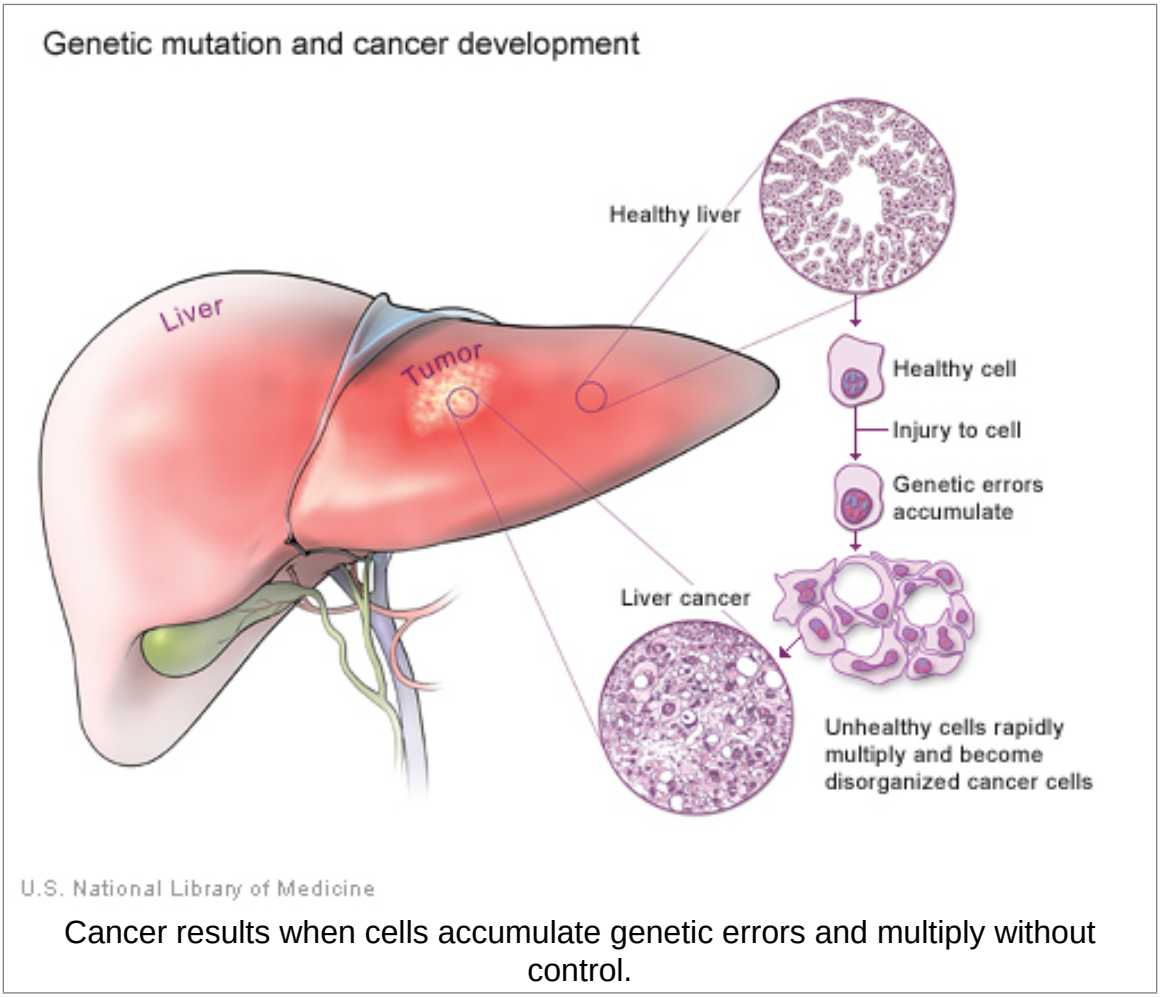
Illustrations



A damaged cell may undergo apoptosis if it is unable to repair genetic errors.



When a cell undergoes apoptosis, white blood cells called macrophages consume cell debris.



Cancer results when cells accumulate genetic errors and multiply without control.

Genetics Home Reference
Your Guide to Understanding Genetic Conditions

http://ghr.nlm.nih.gov/

Lister Hill National Center for Biomedical Communications
U.S. National Library of Medicine
National Institutes of Health
Department of Health & Human Services

Handbook
Help Me Understand Genetics

Chapter	Last Comprehensive Review
The Basics: Genes and How They Work	January 2003

Published on September 22, 2006